BISS Biodiversity Information Science and Standards

Biodiversity Information Science and Standards 9: e174854
doi: 10.3897/biss.9.174854

OPEN ACCESS

Conference Abstract

Reconnecting Indigenous Data to Country

Rose Barrowcliffe‡, Dakota Feirer‡, Alex Ip§, Ben Foley|

‡ Macquarie University, Sydney, Australia
§ AARnet, Canberra, Australia
| University of Queensland, Brisbane, Australia

Corresponding author: Rose Barrowcliffe (rose.barrowcliffe@mq.edu.au)

Received: 14 Oct 2025 | Published: 16 Oct 2025

Citation: Barrowcliffe R, Feirer D, Ip A, Foley B (2025) Reconnecting Indigenous Data to Country. Biodiversity Information Science and Standards 9: e174854. https://doi.org/10.3897/biss.9.174854

Abstract

Herbaria hold millions of specimens collected from Indigenous lands around the world. The specimens may represent a piece of the puzzle that allows Indigenous peoples to care for their Traditional Country ('Country' for short), build sustainable incomes, and continue cultural practices. Indigenous data sovereignty (IDSov) means people have the right to manage and control these specimens as part of their environment and resources (Carroll et al. 2021, Maiam nayri Wingara and Australian Governance Institute 2018), but Indigenous peoples' ability to exercise these rights is limited due to the poor findability of the specimens.

This project, led by a Butchulla and Bundjalung-Gumgganggbir Macquarie University research team and partnering with New York Botanical Garden (NYBG), Local Contexts, the Language Data Commons of Australia (LDaCA), and Australian Academic and Research Network (AARNet), pilots processes for improving the findability of, and reconnecting botanical specimens to Aboriginal and Torres Strait Islander peoples. NYBG staff assisted with understanding the priorities and processes that have resulted in some Australian records being findable but not others. The Local Contexts staff shared learnings from metadata enrichment to similar items in other collections. LDaCA provided initial programming support to map the specimens' collection points and suggested language datasets to indicate who the relevant Traditional Owners might be. AARNet provided the geospatial analysis necessary for the process of inferring connections between large datasets of botanical specimens and Aboriginal groups.

© Barrowcliffe R et al. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Reconnecting the NYBG specimens to their Aboriginal origins involved three steps. First, a dataset of NYBG's Australian-sourced specimens was exported from the Global Biodiversity Information Facility (GBIF). The second step was mapping those specimens to their collection points using the specimens' geolocation metadata. The final step is inferring a relationship between the specimen and Aboriginal Traditional Owner (TO) groups. It is not the place of collecting institutions to decide whose Country is whose; that is for Indigenous people to decide, but in Australia, there is a dearth of Aboriginal-created and owned datasets that self-define Country. As an interim measure, we used open geospatial datasets that have a range of Aboriginal-derived data that describe some forms of connection to Country. These include the Native Lands dataset*[1], the Asher and Moseley Glottography (The Glottography Consortium 2007), and the Native Title Tribunal's Native Title Determinations and Indigenous Land Use Agreements datasets*[2]. Situating the collection points within geographic areas defined in these datasets was used to infer relationships between the NYBG specimens and TO groups.

As an extra measure, we audited a selection of NYBG Australian specimens' metadata in both physical and digital form (through a dataset derived from the NYBG's C.V. Starr Virtual Herbarium online catalogue). This highlighted that there were large amounts of "dark data" in NYBG collections that lack digitisation, processing through optical character recognition (OCR), or transcription to make them findable to Aboriginal peoples through the C.V. Starr Virtual Herbarium and GBIF. Specimens can also be considered dark data if they lack appropriate metadata to make the record findable from a user's point of view. In our pilot, the lack of geolocation metadata inhibited our ability to connect the specimen to Country. While we could not control NYBG's digitisation, OCR and transcription prioritisation, we could enrich specimen geolocation metadata to maximise the number of specimens that could be reconnected to Country. The project plan was amended to include an extra step of using natural language processing (NLP) to parse place names from the 'locality' metadata. The place names were then looked up in the Australian Place Name Composite Gazetteer dataset, resulting in a further 3224 specimens with geolocation metadata.

The final phase of the pilot is to feed the enriched metadata back into the NYBG catalogue to assist Aboriginal peoples in finding the specimens collected from their Country. The ultimate goal of the project is for Aboriginal peoples to work directly with NYBG to exercise their rights in relation to their botanical kin stewarded by NYBG.

Keywords

IDSov, findability, herbaria, botanical specimens, geolocation metadata

Presenting author

Rose Barrowcliffe

Presented at

Living Data 2025

Acknowledgements

The authors of this article acknowledge Indigenous peoples as knowledge and land custodians for many lands around the world for many thousands of years.

Conflicts of interest

The authors have declared that no competing interests exist.

References

- Carroll SR, Herczog E, Hudson M, Russell K, Stall S (2021) Operationalizing the CARE and FAIR Principles for Indigenous data futures. Scientific Data 8 (1). https://doi.org/10.1038/s41597-021-00892-0
- Maiam nayri Wingara, Australian Governance Institute (2018) Indigenous Data Sovereignty Communique. URL: https://static1.squarespace.com/static/5b3043afb40b9d20411f3512/t/5b6c0f9a0e2e725e9cabf4a6/1533808545167/Communique%2B-%2BIndigenous%2BData%2BSovereignty%2BSummit.pdf
- The Glottography Consortium (2007) Glottography Dataset Derived from Asher and Moseley 2007 “Atlas of the World’s Languages”. Zenodo https://doi.org/10.5281/zenodo.15287258

Endnotes

*1 Native Land Digital. 2024. ‘Home’. https://native-land.ca/.

*2 National Native Title Tribunal. n.d. ‘Data Downloads’. Retrieved 13 October 2025. https://www.nntt.gov.au/assistance/Geospatial/Pages/DataDownload.aspx.